Research Article

SYSTEMATIC CONSERVATION PLANNING FOR INTRASPECIFIC GENETIC DIVERSITY.

Ivan PAZ-VINAS[1,2,3*], Géraldine LOOT[1,4], Virgilio HERMOSO[6], Charlotte VEYSSIERE[1,4], Nicolas POULET[7], Gaël GRENOUILLET[1,4] and Simon BLANCHET[1,5]

[1] Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), École Nationale de Formation Agronomique (ENFA); UMR 5174 EDB (Laboratoire Évolution & Diversité Biologique), 118 route de Narbonne, 31062 Toulouse cedex 4, France.

[2] Aix-Marseille Université, CNRS, IRD, Avignon Université; UMR 7263 IMBE, Équipe EGE, Centre Saint-Charles, Case 36, 3 place Victor Hugo, 13331 Marseille cedex 3, France.

[3] Université de Lyon, CNRS, ENTPE; UMR 5023 LEHNA, Laboratoire d'Écologie des Hydrosystèmes Naturels et Anthropisés, 6 rue Raphaël Dubois, 69622 Villeurbanne, France.

[4] Université de Toulouse, UPS; UMR 5174 EDB, 118 route de Narbonne, 31062 Toulouse cedex 4, France.

[5] CNRS, Station d'Écologie Théorique et Expérimentale; UMR 5321, 09200 Moulis, France.

[6] Centre Tecnologic Forestal de Catalunya, Crta. Sant Llorenc de Monunys, Km 2, 25280, Solsona. Lleida, Spain.

[7] Pôle Écohydraulique ONEMA-IMFT, Office National de l'Eau et des Milieux Aquatiques, France.

Running head: Conservation of genetic diversity

Keywords: biodiversity conservation, dendritic ecological networks, multi-specific, freshwater fish, conservation genetics, spatial patterns of biodiversity

***Corresponding author:**

Ivan PAZ-VINAS
Université Lyon 1
CNRS, ENTPE, UMR 5023 LEHNA
Laboratoire d'Écologie des Hydrosystèmes Naturels et Anthropisés
3, rue Raphaël Dubois - Bât. Darwin C
69622 Villeurbanne Cedex France
Phone: (+33) 04 72 43 26 03
E-mail: ivanpaz23@gmail.com

Co-author's e-mails:

Simon Blanchet: simon.blanchet@sete.cnrs.fr
Virgilio Hermoso: virgilio.hermoso@gmail.com
Charlotte Veyssière: veyssiere.charlotte@gmail.com
Nicolas Poulet: nicolas.poulet@onema.fr
Gaël Grenouillet: gael.grenouillet@univ-tlse3.fr
Géraldine Loot: geraldine.loot@univ-tlse3.fr

Abstract

Intraspecific diversity informs the demographic and evolutionary histories of populations, and should hence be a main conservation target. Although approaches exist for identifying relevant biological units of conservation, attempts to identify priority conservation areas for intraspecific diversity are scarce, especially within a multi-specific framework. We used neutral molecular data on six European freshwater cyprinid fish species (*Squalius cephalus*, *Phoxinus phoxinus*, *Barbatula barbatula*, *Gobio occitaniae*, *Leuciscus burdigalensis* and *Parachondrostoma toxostoma*) sampled at the riverscape scale (i.e. the Garonne-Dordogne River basin, France) to determine hot- and cold-spots of genetic diversity, and to identify priority conservation areas using a systematic conservation planning approach. We demonstrate that systematic conservation planning is efficient for identifying priority areas representing a predefined part of the total genetic diversity of a whole landscape. Classical indices of genetic diversity (allelic richness, genetic uniqueness), but private allelic richness, were poor predictors for identifying these priority areas. Moreover, we identified weak congruencies and surrogacies among conservation solutions found for each species, implying that conservation solutions are highly species-specific. Nonetheless, we showed that priority areas identified using intraspecific genetic data from multiple species provide more effective conservation solutions than areas identified for single species or on the basis of traditional taxonomic criteria.

1. Background

Biodiversity conservation is often addressed by identifying protected areas with high biodiversity and/or landscape values [1]. Conservation areas are generally identified as areas with high proportions of endemic, rare or iconic species [2]. Alternatively, conservation planning can be based on the concept of complementarity between conservation areas [3], and on cost-effectiveness analyses such as systematic conservation planning methods (hereafter SCP [4]). SCP aims at identifying a number of complementary and *irreplaceable* sites best representing a predefined proportion of the biodiversity observed in a region and that should be managed for conservation in priority at a minimum cost.

In the meantime, there have been major conceptual and empirical leaps advocating the importance of considering species evolutionary history into conservation plans [5,6]. For instance, there have been attempts to include information on the phylogenetic history of species assemblages into SCP approaches to preserve both species identities and their macro-evolutionary history [7–9]. However, genetic diversity observed at the *population level* (i.e. within species) has rarely been considered in SCP [but see 10–13]. *Intraspecific genetic diversity* is a fundamental facet of biodiversity, as it informs the demographic and colonisation history of populations (i.e. changes in effective population size over time and space), their level of biological connectivity, as well as their potential for adapting to environmental changes [14–17]. Conservation geneticists generally use neutral genetic diversity (i.e. part of the genome not affected by adaptive processes) to identify relevant and significant evolutionary units within a set of individuals sampled in a landscape [18], mainly by analysing their levels of genetic differentiation or (dis)continuity [19,20]. For instance, the “Evolutionary Significant Units” (ESUs) framework has been proposed for representing

different lineages for conservation prioritization [21,22]. By emphasizing complementarity among areas –rather than differentiation–, SCP appears as a promising and complementary framework for identifying an optimal set of sites allowing preserving a pre-defined amount of the total genetic diversity observed within a landscape.

Most SCP studies that have yet considered intraspecific genetic diversity aimed at identifying efficient surrogates for preserving genetic diversity [12,13]. For instance, two recent studies showed that protected areas identified using traditional species distribution data [13], and environmental and geographic descriptors [12] can be efficient surrogates in conservation planning for representing intraspecific genetic diversity. Although these studies shed light on how SCP can help preserving the genetic diversity of multiple species, only few studies have directly used raw genotypic data obtained from multiple species as source of information for SCP analyses (see [11] for the first attempt using data from a single species). Datasets generated using neutral genetic markers such as microsatellite markers (i.e. alleles presence/absence matrices) are similar to datasets generally used for SCP (i.e. species presence/absence matrices), hence making possible applying SCP to genotypic surveys without needing intermediate modelling procedures or the calculation of summary statistics of genetic diversity (as in [10,12] for instance) that may fail to capture the whole genetic diversity of populations.

Here, we tested the potential of SCP analyses to identify priority conservation areas accounting for intraspecific genetic diversity measured for a species assemblage at the landscape scale (i.e. at the scale of an entire river basin). We first considered a set of four common and representative freshwater fish species (*Squalius cephalus*, *Gobio occitaniae*, *Barbatula barbatula* and *Phoxinus phoxinus*) to test the influence of

conservation targets (proportion of the total amount of genetic diversity to be secured by irreplaceable sites) and analytical strategies (analysing each species independently or all species pooled) on final conservation solutions (number and identity of irreplaceable sites). We then included two rare species of particular conservation interest (*Leuciscus burdigalensis* and *Parachondrostoma toxostoma*) to test the relevance of the SCP approach in a "real conservation-oriented study". For these two species (and the four common species), we ran SCP analyses considering a typical conservation target [23,24] to explore the spatial distribution of irreplaceable sites in a riverscape, and to test for congruence and surrogacy in irreplaceable sites among species, and more particularly between rare and common species. We further tested whether irreplaceable sites identified using SCP analyses are more efficient for preserving intraspecific genetic diversity than sites randomly selected or than sites selected on the basis of traditional taxonomic criteria (e.g. sites displaying high species richness). We finally tested whether irreplaceable sites were correctly predicted by classical indices of genetic diversity such as allelic richness. We demonstrate that preserving the intraspecific genetic diversity of a species assemblage using systematic conservation planning is an efficient and relevant, yet complex, approach necessitating appropriate analyses to assist the decision-making process.

2. Methods

<u>Data collection</u>

Biological models. We focused on an assemblage of six cyprinid freshwater fish species. Three of them are widespread in Europe (i.e. *Squalius cephalus*, *Phoxinus*

phoxinus and *Barbatula barbatula*) whereas three of them (*Gobio occitaniae*, *Leuciscus burdigalensis* and *Parachondrostoma toxostoma*) are endemic to Southern France [25]. This set of species covers a large functional trait space that is representative of many freshwater fish communities. For instance, *S. cephalus* is a large-bodied fish species with a long lifespan (i.e. it can be 60 cm long and live up to 15 years [25]) whereas at the other extreme *Ph. phoxinus* is a small-bodied species with a shorter lifespan (i.e. it is less than 12 cm long and usually lives up to 4-5 years [25]). From an ecological perspective, *G. occitaniae, Pa. toxostoma* and *B. barbatula* are bottom feeders, whereas *S. cephalus* and *Ph. phoxinus* are water column feeders and *L. burdigalensis* is more opportunistic. Further, *B. barbatula* is mainly active during the night, while the other species are particularly active during the day. Four of these species are relatively abundant (i.e. *S. cephalus*, *G. occitaniae*, *Ph. phoxinus* and *B. barbatula*), although they vary greatly in their ecological niche and hence their spatial occupancy in the river network (see Figure S1 for maps representing the spatial distribution of sampling sites for each species, which roughly corresponds to their spatial distribution in the Garonne-Dordogne river basin). We will hereafter refer to this set of species as the "common" species. The two other species are rare (*L. burdigalensis*) to very rare (*Pa. toxostoma;* see Figure S1) in the Garonne-Dordogne river basin, and are of particular interest for conservation. *Leuciscus burdigalensis* is a recently described species that is locally experiencing both demographic and genetic bottlenecks in many populations [26,27]. *Parachondrostoma toxostoma* is a vulnerable species [28] listed in the IUCN red list, in the Annex II of the European Union Habitats Directive and in Appendix III of the Bern Convention [28].

Sampling design. During Spring/Summer 2010-2011, we used electrofishing to sample 92 sites distributed across 35 rivers from a whole European river basin, the Garonne-Dordogne River basin (>100,000 km^2, South-Western France; Figure S1; Table S1). Sampling sites were chosen to cover the entire distribution of each species at the riverscape scale, and to allow characterising spatial patterns of genetic diversity for all these species. Up to 25 individuals *per* species *per* site were sampled when possible. Not all species were present at all sampling sites (Figure S1; Table S1), and not all sites provided 25 individuals *per* species due to low densities. In these cases, we captured as many individuals as possible. We anesthetized each individual and then we collected and stored in 90% ethanol a fragment of their pelvic fin. All individuals were released alive at their sampling location.

Genotyping. Genomic DNA was extracted using a salt-extraction protocol [29]. We used multiplexed Polymerase Chain Reactions (PCRs) to co-amplify 8 to 15 microsatellite loci depending on the species (8 for *G. occitaniae*, 9 for *B. barbatula*, 10 for *S. cephalus* and *Ph. phoxinus*, 14 for *L. burdigalensis* and 15 for *Pa. toxostoma*). We used 5-20 ng of genomic DNA and QIAGEN® Multiplex PCR Kits (Qiagen, Valencia, CA, USA) to perform PCR amplifications. Details on loci, primer concentrations, PCR conditions and multiplex sets can be found in Table S2. The genotyping was conducted on an ABI PRISM™ 3730 Automated Capillary Sequencer (Applied Biosystems, Foster City, CA, USA). The scoring of allele sizes was done using GENEMAPPER® v.4.0 (Applied Biosystems).

Genetic diversity assessment. Given the good spatial resolution of the sampling obtained for the four common species (i.e. *S. cephalus*, *G. occitaniae*, *Ph. phoxinus* and *B.*

barbatula), descriptive genetic analyses were conducted for sampling sites displaying a minimum sample size of N=10 individuals for these species, so as to maximize consistency on subsequent allelic frequency-based genetic analyses (see Figure S1 and Table S1 for details on sample sizes). For the two rare species, for which the sampling was more restricted (i.e. *L. burdigalensis* and *Pa. toxostoma*), genetic analyses were conducted for sampling sites displaying a minimum sample size of N=6 individuals, so as to maximize the number of sampling sites included in the SCP procedures. This allowed us to considerably increment the number of sites retained for genetic analyses by 38.1% (21 to 29 sites) and 75% (from 8 to 14 sites) for *L. burdigalensis* and *Pa. toxostoma* respectively and hence to improve the representation of the spatial distribution of these two species. We then determined for each of the six species the occurrence of null alleles and potential scoring errors with the program MICROCHECKER 2.3 [30]. We tested for departures from Hardy-Weinberg (HW) equilibrium with the 'adegenet' R package v1.6-2 [31]. The program GENEPOP v4.0 [32] was used to assess linkage disequilibrium among loci within sites. Levels of significance for these multiple tests were adjusted using Bonferroni procedures. We found significant deviations from HW for a few locus/population pairs for each of the six species (see Appendix A1 and Supplementary Data File 1 for details and raw tables), and significant linkage disequilibrium and homozygote excesses only for the four common species (Appendix A1; Supplementary Data File 1). However, no clear patterns were observed for any species across loci and populations for these deviations. Given the small extent of these deviations and the large spatial extent of the databases, we assumed that they weakly affected our main findings (Appendix A1; Supplementary Data File 1).

To assess within-site intraspecific genetic diversity (i.e. α-IGD), we applied rarefaction procedures implemented in ADZE v1.0 [33] to calculate both allelic richness (*AR* [34]) and private allelic richness (*PA* [35]) at the sampling site level (based on a minimum of N=10 individuals for common species or N=6 individuals for rare species). To assess the among-site component of intraspecific genetic diversity (i.e. β-IGD), we used the R package 'mmod' [36] to calculate –for each species– a pairwise genetic differentiation index (i.e. D_{est} [37]). For each site (and species), we then derived the averaged value of all pairwise D_{est} values estimated between one given site and all the remaining sites, so as to obtain a single value *per* site. For rare species, the impact of considering sites with N≤6 individuals instead of N≤10 on *AR*, *PA* and D_{est} was relatively low (i.e. no clear impact neither on *PA* nor on D_{est} values, and a mean decrease in *AR* values of 0.295 alleles *per* site for *Pa. toxostoma* and of 1.183 alleles *per* site for *L. burdigalensis*). This impact was however overcompensated by the increase in the number of sites retained for these two species when considering sites with N≤6 individuals.

<u>Testing the suitability of SCP analyses for intraspecific genetic diversity</u>

In the first step, we tested the influence of conservation targets and analytical strategies on final conservation solutions. We focused specifically on data from the four common species, as their large coverage of the sampling area is more suited for the demonstrative exercise done in this step.

Spatial patterns of genetic diversity. We first used geostatistical modelling tools to explore spatial patterns of α and β genetic diversity for the four common species at the

riverscape scale by predicting the distributions of *AR*, *PA* and D_{est} from the observed empirical values using Generalized Linear Models for Spatial Stream Networks (GLMSSN [38,39]). This was done using the 'STARS' toolset of ARCGIS v10.2 and the R package 'SSN' [38,39]. We conducted a model selection procedure based on a comparison of Akaike Information Criteria (AIC) estimated for several competing GLMSSNs. These models were built by (i) assuming three geographic descriptors (i.e. topological distance from the outlet, longitude and latitude) as explanatory variables, and (ii) choosing a tail-down covariance structure model among the following ones: exponential, Mariah, spherical, linear-with-sill and Epanechnikov. The model with the lowest AIC score (see Supplementary Data File 1 for raw results) was considered as the best model and was used to estimate the slope and the significance of the relationships between explanatory variables and each genetic index, which allowed testing whether or not spatial patterns of intraspecific genetic diversity can be detected. We finally used predictions from the best models to produce krigged maps for each common species and each genetic index, so as to visually represent the spatial distribution of intraspecific genetic diversity across the whole river drainage, and to visually highlight hot- and cold-spots of intraspecific genetic diversity. We also calculated Pearson's correlation coefficients between values calculated at the site level for each genetic index (i.e. *AR*, *PA* and D_{est}) for each pair of common species, so as to test for spatial congruence in patterns of genetic diversity among common species.

Identification of irreplaceable sites. We then tested whether conservation targets (i.e. the percentage of total number of conservation units to be present in the final conservation solution) and analytical approaches (i.e. species-specific or species-pooled analyses) influence the identification of irreplaceable sites. SCP methods traditionally

use species presence/absence data as input data to identify irreplaceable sites for the conservation of taxonomic diversity at the community level (i.e. sites that cannot be excluded from an optimal selection of sites for conservation [40]). Here, we replaced species presence/absence data with allele presence-absence data as in [11] to identify irreplaceable sites for the conservation of intraspecific genetic diversity of each species in the river drainage. We used the program Marxan v2.1 [40] and genotypic data from the common species to identify, for each common species independently, an optimal set of sites that best represent at least 50, 75, 90 or 100% of the total number of alleles present in the whole riverscape at a minimum "cost", which corresponds to four conservation targets. Given the lack of ground estimates for conservation cost, we used a constant cost *per* site, so our objective translated into identifying the minimum number of sites that represent a given proportion of intraspecific genetic diversity (i.e. 50, 75, 90 or 100% of the total pool of alleles represented at least once [41]). We arbitrarily choose the 50, 75 and 100% conservation targets to explore how the proportion of alleles to protect affects the selection of irreplaceable sites. We additionally tested the 90% conservation target, as it corresponds to a threshold target being typically assumed in *ex-situ* conservation plans [23,24]. In order to estimate the relevance of each site to preserve a given proportion of the allelic diversity in the river basin, we used two different methods. In case of 100% of allelic diversity, we used the traditional irreplaceability measure reported by Marxan that ranges between 0% (not irreplaceable) and 100% (highly irreplaceable). This measure is estimated by running the optimization algorithm a number of times (N=100 runs in our case) and then computing the frequency of selection of each site within the solutions obtained. Sites with unique allelic composition will be selected across all runs and reported as highly irreplaceable, whereas sites with more common alleles, replaceable by other sites with

the same alleles, will appear poorly irreplaceable. For the other conservation targets, we selected a random pull of 50, 75 and 90% of the total number of alleles existing at the basin level and for each species. We then ran the optimization algorithm to identify the minimum set of sites representing this particular selection of alleles. Given that the analyses were run only on a subset of alleles, we replicated this process 100 times to minimize the effect of the arbitrary selection of alleles. For each subset of alleles, we ran the Marxan procedure as explained above (e.g. constant cost and 100 runs for each) and retained the best solution for subsequent analyses (solution with the lowest value for the objective function). We then calculated the irreplaceability as the frequency of selection of each site within the 100 random pull of alleles. For a given species and a given conservation target, we considered a site as irreplaceable for genetic diversity conservation when it displayed an irreplaceability value of 100%. We selected such a high threshold so as to be conservative. We tested and compared visually how the proportions of irreplaceable sites vary among species and conservation targets.

To test how pooling data from several species affect the identification of irreplaceable sites, we further performed a “pooled” analysis, in which all alleles found for each common species at a site were pooled together in a single input dataset. We then selected a random pull of 30, 50, 75, 90 and 100% of the total number of alleles existing at the basin level (all common species confounded), and performed 100 Marxan runs *per* conservation targets to identify the minimum set of sites representing these particular selections of alleles. Allele counts for all species are provided in the Supplementary Data File 1.

<u>A real conservation-oriented study using SCP approaches</u>

In this second step, we (i) explored the spatial distribution of irreplaceable sites, (ii) tested whether conservation solutions found for each species are congruent among species, (iii) tested whether solutions found for one species can be used as a surrogate for other species, (iv) assessed the overall efficiency of the different conservation solutions we identified for preserving the overall genetic diversity of all species (v) compared the efficiency of conservation solutions identified through SCP to the efficiency of alternative conservation solutions based on traditional taxonomic criteria (e.g. in sites with high species richness), and finally (vi) tested whether –or not– irreplaceable sites are correctly predicted by classical indices of genetic diversity. In this step, we included the two rare species, since congruence and surrogacy are particularly important to measure for rare species for which data are more difficult to collect. We therefore focus more specifically on the comparisons comparing common *vs*. rare species.

Identification of irreplaceable sites for genetic conservation. We focused only on irreplaceable sites identified for the 90% target and used the program Marxan as described above to identify these sites for *L. burdigalensis* and *Pa. toxostoma* independently, in addition of the four common species.

We mapped these irreplaceable sites (for the six species pooled or independently) on the river network, so as to test (i) whether or not specific areas harboured more irreplaceable sites (e.g. upstream areas that are generally thought to be of high conservation priority [42]) and (ii) if irreplaceable sites are spatially congruent among species and, most notably, among common and rare species. In addition, we ran GLMs (assuming a binomial error terms distribution) including whether or not a site has been designated as an irreplaceable site at the 90% target as a binomial dependent variable,

and distance to the outlet of sampling sites and betweenness centrality values [43,44] for each sites explanatory variables. Betweenness centrality is an index quantifying the positional importance of each sampling site within the river basin [45,46]. The significance of each term was tested at the α=0.05 threshold.

Surrogacy in irreplaceable sites among species. We then estimated the levels of surrogacy among species by calculating the percentage of the total number of alleles observed for a given species that is secured by irreplaceable sites identified for another species. Although surrogacy was calculated for all species pairs, we specifically focused on rare species by calculating (i) the percentage of the total number of alleles observed for rare species secured by all the irreplaceable sites identified for all the common species, and (ii) the percentage of the total number of alleles observed for common species secured by all the irreplaceable sites identified for the rare species.

Assessing the efficiency of SPC. The goal of this analysis was (i) to assess the efficiency of SCP for preserving the overall levels of intraspecific genetic diversity observed at the riverscape scale for the six species, and to (ii) compare the efficiency of SCP-based conservation solutions to the efficiency of alternative conservation solutions selected on the basis of traditional taxonomic criteria. We first computed an average accumulation curve of “protected genetic diversity” by estimating the average cumulative percentage of the total number of alleles (over all six species) secured by a cumulative number of randomly selected sampling sites. Specifically, we increased the number of sites being sampled from 1 to 92 with a continuous increment of one unit. At each incremental step, we randomly selected sampling sites (without replacement) and we calculated the percentage of the total number of alleles (over all species) secured by the random set of

sites. We repeated this random selection 1000 times for each incremental step and averaged the percentages over these 1000 permutations to obtain the averaged accumulation curve of protected genetic diversity. We then tested to which extent, for an identical number of selected sites, the different conservation solutions we identified through SCP (i.e. irreplaceable sites) secured higher percentages of genetic diversity than those secured by a random selection of sites. Notably, we explored whether certain conservation solutions (i.e. by considering irreplaceable sites identified for single species or for the sets of common species, rare species and all the six species) were more efficient than others to sustain a large percentage of the total number of alleles.

As in most conservation studies species diversity is the main conservation target [47], we finally checked the efficiency of conservation solutions based on taxonomic criteria for preserving the overall levels of genetic diversity of the six species of interest. Specifically, we calculated the percentages of total number of alleles (over all species) secured by sites (i) in which a number equal or greater than 10, 15 or 20 freshwater fish species co-occur (the total number of fish species observed in the riverscape being 34 and the mean number of species observed at the site level being 12.77 +/- 0.55), and (ii) in which the six fish species considered in our study co-occur. We finally compared the efficiency of taxonomic-data-based conservation solutions to that of SCP-based conservation solutions to assess if solutions based on taxonomic criteria are good surrogates of the overall intraspecific genetic diversity of our species assemblage. Data on fish species occurrence at the site level has been derived from [46] for 81 out of our 92 sampling sites.

Relationships between irreplaceable sites and indices of genetic diversity. To test the ability of classical genetic indices to predict the propensity for a site to be irreplaceable

from a conservation viewpoint, we ran for each species Generalized Linear Models (GLMs, with a binomial error terms distribution) including whether or not a site has been designated as an irreplaceable site at the 90% target as a binomial dependent variable, and *AR*, *PA* and D_{est} as explanatory variables. We tested the significance of each term at the α=0.05 threshold. Explanatory variables were centred and scaled, in order to compare the relative strength of the predictors among species.

3. Results

Descriptive statistics. Overall, *Pa. toxostoma* (one of the two rare species) showed the lowest within-site intraspecific genetic diversity estimates. Mean *AR* ranged from 2.114 for *Pa. toxostoma* to 5.821 for *Ph. phoxinus*, and mean *PA* ranged from 0.036 for *Pa. toxostoma* to 0.162 for *L. burdigalensis* (Table S3). *Parachondrostoma toxostoma* also showed the lowest mean D_{est} value (0.069), while the highest mean value was found for *B. barbatula* (0.383; Table S3).

<u>Testing the suitability of SCP analyses for intraspecific genetic diversity</u>

Spatial patterns of genetic diversity. Spatial patterns of genetic diversity largely varied and were actually poorly congruent among the four common fish species (Figure 1). As extreme examples, hotspots of *AR* for *S. cephalus* where mainly found in the Western part of the riverscape and on the core streams, whereas these same areas were identified as coldspots of *AR* for *B. barbatula* (Figure 1A1-1A3). Similarly, hotspots of *PA* were inversely related between *G. occitaniae* and *Ph. phoxinus* (Figure 1B2-1B4). Similar conclusions were reached for D_{est} (Figure 1C). For instance, hotspots of D_{est} were observed in opposite areas of the riverscape for the species pair *B. barbatula/Ph.*

phoxinus (Figure 1C3-1C4). As a consequence, the sign, slope and significance of GLMSSNs explanatory variables strongly varied among species (Figure 1). This incongruence in spatial patterns of genetic diversity among species was also reflected by the low to moderate correlation coefficients measured among all possible pairs of species and for each index of genetic diversity (i.e. *AR*, *PA* and D_{est}, Table S4). Indeed, Pearson's correlation coefficients were lower than 0.6 for all comparisons but two (i.e. between *B. barbatula*/*G. occitaniae* and between *Ph. phoxinus*/*B. barbatula* for *AR*, see Table S4).

The influence of conservation targets and analytical strategies to identify irreplaceable sites for genetic conservation. When species were analysed independently, we found that the number of irreplaceable sites increased as the conservation target (i.e. the percentage of alleles secured by irreplaceable sites) increased, with a steep increase for conservation targets higher than 75% of the total number of alleles (Figure 2). However, the percentage of irreplaceable sites strongly varied among species. For instance, for a 90% conservation target, the proportion of irreplaceable sites ranged from 3.61% of the total number of sampled sites for *G. occitaniae* to 28.57% for *Ph. phoxinus* (Table S5; Figure 2). For extreme conservation targets (100% of alleles to be secured), the proportion of irreplaceable sites varied from 25.30% for *G. occitaniae* to 68.26% for *Ph. phoxinus* (Table S5; Figure 2).

When alleles from the four common species were analysed in a single pooled analysis, we similarly found that the proportion of irreplaceable sites increase as the conservation target increases (Figure 2). Interestingly, we did not identify irreplaceable sites for the 30% conservation goal, and only 3 irreplaceable sites were found for the 50% conservation goal (Figure 2). The proportion of irreplaceable sites increased

moderately to 17.39% for the 75% target, and then steeply increased for higher conservation targets to reach 55.70% for the 90% target and 76.08% for the 100% target (Figure 2). This later result suggests that almost all the riverscape should be protected to reach high conservation targets when adopting a pooled strategy.

A real conservation-oriented study using SCP approaches

Identification of irreplaceable sites for genetic conservation. We first visually explored the spatial distribution of irreplaceable sites. Overall, the localization of irreplaceable sites in the riverscape strongly varied among species (Figure 3). We failed to identify areas (e.g. upstream or downstream locations) clustering irreplaceable sites for any species (Figure 3A-F). This apparent lack of clustering was statistically confirmed by our generalized linear models (Table S6). Indeed, neither the distance from the outlet of sampling sites nor the betweenness centrality of sampling sites were significant predictors of the irreplaceability of sites for all species (the only exception being a significant effect of distance from the outlet for *Ph. phoxinus*, Table S6).

Second, we tested whether conservation solutions found for each species were spatially congruent among species. Over the six fish species, we identified forty-two sites (45.65%) as irreplaceable at the 90% conservation target for at least one species (Figure 4). Thirty-two of these forty-two sites were irreplaceable for at least one of the four common species (Figure 4), and fourteen of the forty-two sites were irreplaceable for at least one of the two rare species (Figure 4). Among the six species, only eight out of these forty-two sites were irreplaceable for at least two species, and only one of these sites was irreplaceable for three species (Figure 4).

Surrogacy in irreplaceable sites among species. Levels of surrogacy were generally low to moderate, and strongly varied among species pairs (Table 1). For instance, irreplaceable sites identified for *G. occitaniae* failed to secure the genetic diversity of the two rare species, and covered only 18 to 43% of the total number of alleles of the other common species. This indicates that irreplaceable sites identified for *G. occitaniae* (i.e. the most widespread species) are poor surrogates for preserving the intraspecific genetic diversity of other species (18.38% of surrogacy; Table 1). Conversely, irreplaceable sites found for *Ph. phoxinus* and *L. burdigalensis* are better surrogates for *G. occitaniae*, as 79.58% and 70.27% of the total number of alleles of *G. occitaniae* was secured by irreplaceable sites identified for *Ph. phoxinus* and *L. burdigalensis* respectively (Table 1). Overall, irreplaceable sites best sustaining genetic diversity of other species were those identified for *B. barbatula*, which secured in average 68.28% of the total number of alleles of other species (Table 1).

Interestingly, the thirty-two irreplaceable sites identified for the four common species secured 81.79% and 90% of the total number of alleles of *L. burdigalensis* and *Pa. toxostoma* respectively, suggesting that irreplaceable sites identified for a set of common species can be good surrogates for intraspecific genetic diversity of rare species. The fourteen irreplaceable sites identified for the two rare species secured a total number of alleles ranging from 65.18% for *S. cephalus* to 79.40% for *B. barbatula*, which is also not negligible given that conservation priorities are generally directed towards rare species.

Assessing the efficiency of SPC. Overall, irreplaceable sites identified for single species using SCP did not significantly secured higher percentages of genetic diversity than randomly selected sites, excepting irreplaceable sites identified for *L. burdigalensis* and

B. barbatula (Figure 5), which were more efficient for preserving the overall genetic diversity of all species than randomly selected sites. However, we found significantly higher percentages of secured alleles than randomly selected sites on irreplaceable sites identified through SCP for (i) the set of common species (ii) the set of rare species, and (iii) all the six species (Figure 5).

Conservation solutions based on the co-occurrence of high numbers of species (i.e. sites with numbers of co-occurring species equal or greater than 10, 15 or 20) did not secured more alleles than those found for randomly selected sites, suggesting that raw interspecific diversity (i.e. species richness) may not be a good surrogate of intraspecific genetic diversity in our study (Figure 5). Nevertheless, we found that a higher percentage of alleles would be secured by selecting sites on which the six species co-occur than by randomly selecting sites in the riverscape (Figure 5).

Relationships between genetic indices and irreplaceable sites. For all species but *G. occitaniae* and *Pa. toxostoma* (for which none of the variables were significant predictors), *PA* was identified as the only variable that significantly predicted the probability for a site to be identified as an irreplaceable site (Table S7). This probability increased with the number of *PA* in a site.

4. Discussion

Intraspecific genetic diversity reflects the evolutionary and demographic history of populations, and hence sustain the capacity of species and populations to cope with environmental changes [14–17]. This biodiversity facet is the first that should respond to global change [16], and is hence surprising that this biodiversity facet has yet rarely

integrated in dedicated optimization planning tools. We here fill this gap by demonstrating that systematic conservation planning for intraspecific genetic diversity is an efficient, yet complex, approach necessitating careful considerations.

From idiosyncratic distributions of genetic diversity…

Our results strongly suggest that the genetic diversity of the targeted species did not follow a common spatial pattern, but rather species-specific (idiosyncratic) spatial distributions. This conclusion holds true for all genetic diversity indices, and it corroborates the few previous studies investigating simultaneously the spatial distribution of genetic diversity at large spatial scales and for sympatric species (e.g. [48,49]). This was however unexpected given that recent meta-analyses on freshwater organisms demonstrated that α-IGD is generally higher in downstream than in upstream areas [50], and that β-IGD tends to be higher in upstream than in downstream sections [42]. Overall, these very general spatial patterns of intraspecific genetic diversity were verified in our datasets. For instance, a negative relationship between allelic richness and distance from the outlet is expected in freshwater organisms [50], and was actually observed for three out of the six species considered in this study (Figure S2). However, when using a precise and novel approach to map genetic diversity across the network, we demonstrated that the distribution of cold- and hot-spots of α- and β-IGD was subtler and idiosyncratic. This probably reflects interactions between recent colonization histories, life-history-traits of species and the network structure, which are expected to drive patterns of neutral genetic diversity in rivers [50–52].

…to systematic conservation planning for intraspecific genetic diversity.

The spatial mismatch in intraspecific genetic diversity among species probably explains why we found that the level of congruence and surrogacy of irreplaceable sites identified for each species was extremely low. For instance, there was an extremely low proportion of irreplaceable sites that were common to two or three species (and never more than three species; Figure 4). In the same way, we detected no clear patterns in the spatial distribution of irreplaceable sites. Indeed, irreplaceable sites (for any of the six species) were not particularly situated in specific locations such as upstream areas and/or in areas of high connectivity (i.e. with high centrality values [45,46]), hence suggesting that priority areas for the conservation of intraspecific genetic diversity of species assemblages should not be restricted to specific areas in the riverscape. It is however noteworthy that the greatly variable position of irreplaceable sites in the riverscape found for the species we studied may reflect their own ecological niche and spatial occupancy in the riverscape. Finally, the level of surrogacy among irreplaceable sites was low to moderate, and never attained the 90% threshold we assumed when we considered all species independently. Combined with our finding that the number of irreplaceable sites can be relatively high for reasonable conservation targets (up to 46% sites are identified as irreplaceable sites for at least one species at the conservation target of 90%), we conclude that our ability to identify priority areas for intraspecific genetic diversity may be highly species-specific and may depend on the capacity to tackle the trade-off between the amount of genetic diversity to protect, and the extent of priority areas we can realistically protect.

However, when surrogacy between all irreplaceable sites identified for the entire set of common species and those identified for the rare species was tested, we reached reasonable proportions of the total number of alleles to protect (~80-90%) for the two rare species (i.e. *Pa. toxostoma* and *L. burdigalensis*). This result suggests that, in some

cases, genetic data obtained for a set of widely-distributed, “easier-to-sample” common species displaying varying life-history traits can be used for identifying protection areas for the intraspecific genetic diversity of other rare species that can be more problematic to sample. This would require, however, a good overlap between the spatial distributions of the common species used as surrogates and the spatial distributions of the targeted rare species, as well as a wide sampling design allowing representing the whole spatial occupancy of species in the riverscape. Failing to fulfil these two prerequisites may have a negative impact on conservation costs, as the number of irreplaceable sites necessary to preserve a predefined amount of genetic diversity of rare species would be higher than expected. In our demonstrative study for instance, thirty-four of the forty-two (81%) irreplaceable sites were sites in which at least four study species co-occur, and nineteen of these sites (54.3%) were sites in which the six species co-occur. Another important finding we report here is that selecting conservation areas based on traditional taxonomic criteria (i.e. by selecting sites with high species richness) is not an efficient strategy for preserving the overall genetic diversity of a set of species. This highlights the need for finding efficient ways for considering intraspecific genetic diversity in systematic conservation planning, as we and others [e.g. 9–13] proposed.

Overall, our results suggest that two different analytical strategies can be employed in real-case SCP studies aiming at preserving intraspecific genetic diversity of a species assemblage (i) identification of conservation areas for a set of rare species or (ii) identification of conservation areas for a set of representative common species. These two strategies have their own advantages and inconveniences. The first strategy may optimally preserve genetic diversity of rare species at competitive costs (e.g. only 14 irreplaceable sites to protect in our demonstrative study), but this at the expense of the genetic diversity of other sympatric species. Conversely, the second strategy may

optimally preserve genetic diversity of a set of common species while maintaining high levels of genetic diversity for rare species, but this at a higher cost (e.g. 32 irreplaceable sites to protect in our study). Whether to choose one of these two strategies will therefore depend on many factors such as how difficult is the sampling of rare species compared to common species, or the extent of the resources available for setting new protected areas. We recommend adopting the second strategy when possible, since it allows to simultaneously maintain genetic diversity from rare and common species. Indeed, genetic diversity of common species is vital for ensuring ecosystem stability, as it ultimately influences species interactions, population dynamics and ecosystem functioning [53], and we argue that it should be considered in conservation plans. It is noteworthy that further studies must be conducted to test if the two analytical strategies we propose can be applicable across taxa and landscapes.

Although it was not a main goal of the study, we identified an alternative strategy that may also be efficient for maintaining high levels of genetic diversity of a species assemblage without requiring SCP analyses. Indeed, selecting sites where all the studied species co-occur as conservation solutions allowed maintaining high levels of genetic diversity for both common and rare species in our study, although this strategy was less efficient than selecting irreplaceable sites found for common species through SCP. The efficiency of this strategy also needs to be thoroughly tested across different taxa and landscapes, as it can be a useful surrogate strategy in some cases (e.g. when molecular data cannot be produced due to funding or technical limitations). It is however noteworthy that such strategy requires highly reliable taxonomic data gathered over the whole studied landscape.

Finally, we demonstrate that microsatellite-based genotypic data can be easily used as direct input data in SCP to identify efficient conservation areas for intraspecific

genetic diversity of a species assemblage at the landscape scale. Relatively small numbers of microsatellite markers such as those considered in our study (8 to 15) have been proven to be sufficient for capturing the confounding genetic effects of recent past demographic history, genetic drift and connectivity issues between populations [27,50,51], and their efficiency for informing recent evolutionary processes can be equal or higher than low to moderate numbers of SNP markers for instance (e.g. <300 SNPs; [19]). However, microsatellite markers have the disadvantage of not being suitable for drawing conclusions on the adaptive potential of populations, contrarily to other markers derived at the genome scale. Given that our capacity to compile genetic datasets at large spatial, temporal and taxonomic scales has greatly increased in recent decades [54], a next methodological step may be the use of high-throughput techniques to develop high numbers of markers informing both neutral and non-neutral processes to assess whether SCP solutions obtained using different categories of markers allow efficiently preserving both the whole genomic diversity of populations, and hence their whole evolutionary history.

5. Conclusions

Our study provides novel, insightful and promising knowledge on the setting of priority conservation areas for intraspecific diversity. It shows that systematic conservation planning methods are useful and efficient objective tools for conservation geneticists whose conservation solutions will strikingly depend on the species to be conserved and the quantity of genetic information that managers aim at preserving in a landscape. Given our results, we suggest that two strategies could be employed in real-case conservation programs: (i) identification of priority conservation areas for a set of

rare species when resources allocated to conservation are scarce or (ii) identification of priority conservation areas on the basis of the analysis of a set of representative common species that may serve as “umbrellas” for rare sympatric species when resources allocated to conservation are higher.

Our study also raises many additional questions that should be considered in the near future. Among others, we believe that the next steps will be to formally identify sound conservation targets for intraspecific diversity, to test whether neutral intraspecific diversity appropriately mirrors quantitative and adaptive diversity [17], and to quantify the influence of intraspecific diversity on ecosystem functioning and services, so as to better evaluate the added value of preserving such a facet of biodiversity [17,55].

Competing Interests statement

Authors declare not having any competing interest.

Author Contributions

The study was designed by SB, GL and IPV. IPV and SB wrote the manuscript with the help of VH, GG, NP and GL. IPV, SB, GL and CV collected the samples. GL and CV produced the genetic data. IPV, VH and SB conducted the population genetic and statistical analyses.

Acknowledgements

We are very grateful to all the ONEMA staff that contributed to the field sampling, in particular Laurence Blanc. Leslie Faggiano, Charlotte Evangelista, Christine Lauzeral, Loïc Tudesque, Roselyne Étienne, Sébastien Villéger, Julien Cucherousset, Vincent

Dubut, Alain Blanchet and South-Western France ONEMA staff are thanked for their help in the field. Roselyne Étienne and Vincent Dubut are also thanked for their help with molecular analyses. Lisa Fourtune is acknowledged for having calculated the betweenness centrality values of our river network. We also thank Marie-Hélène Lizée for her help with GIS tools and Camille Pagès and Jérôme Prunier for their comments. Further, the authors thank the "Génopole Toulouse" for help with genotyping and Keoni Saint-Pée for correcting the English. This work has been done in two research units (EDB & SETE) that are part of the "Laboratoire d'Excellence" (LABEX) entitled TULIP (ANR-10-LABX-41).

Funding

IPV was financially supported by a MESR ("Ministère de l'Enseignement Supérieur et de la Recherche") PhD scholarship during this study. VH was funded by the National Environmental Research Program Northern Australia Hub and a "Ramon y Cajal" contract (RYC-2013-13979) funded by the Spanish government. This study is part of the European project "IMPACT". This project has been carried out with financial support from the Commission of the European Communities, specific RTD programme "IWRMNET".

References

1. Watson JEM, Dudley N, Segan DB, Hockings M. 2014 The performance and potential of protected areas. *Nature* **515**, 67–73.

2. Filipe AF, Marques TA, Tiago P, Ribeiro F, Da Costa LM, Cowx IG, Collares-Pereira MJ. 2004 Selection of priority areas for fish conservation in Guadiana river basin, Iberian Peninsula. *Conserv. Biol.* **18**, 189–200.

3. Kirkpatrick JB. 1983 An iterative method for establishing priorities for the selection of nature reserves: An example from Tasmania. *Biol. Conserv.* **25**, 127–134.

4. Margules CR, Pressey RL. 2000 Systematic conservation planning. *Nature* **405**, 243–253.

5. Vane-Wright RI, Humphries CJ, Williams PH. 1991 What to protect?—Systematics and the agony of choice. *Biol. Conserv.* **55**, 235–254.

6. Faith DP. 1992 Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **61**, 1–10.

7. Asmyhr MG, Linke S, Hose G, Nipperess DA. 2014 Systematic conservation planning for groundwater ecosystems using phylogenetic diversity. *PLoS ONE* **9**, e115132.

8. Buerki S, Callmander MW, Bachman S, Moat J, Labat J-N, Forest F. 2015 Incorporating evolutionary history into conservation planning in biodiversity hotspots. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140014–20140014.

9. Carvalho SB, Velo-Antón G, Tarroso P, Portela AP, Barata M, Carranza S, Moritz C, Possingham HP. 2017 Spatial conservation prioritization of biodiversity spanning the evolutionary continuum. *Nat. Ecol. Evol.* **1**, 0151.

10. Thomassen HA *et al.* 2011 Mapping evolutionary process: a multi-taxa approach to conservation prioritization. *Evol. Appl.* **4**, 397–413.

11. Diniz-Filho JAF *et al.* 2012 Planning for optimal conservation of geographical genetic variability within species. *Conserv. Genet.* **13**, 1085–1093.

12. Hanson JO, Rhodes JR, Riginos C, Fuller RA. 2017 Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proc. Natl. Acad. Sci.*, 201711009.

13. Hermoso V, Kennard MJ, Schmidt DJ, Bond N, Huey JA, Mondol RK, Jamandre BW, Hughes JM. 2016 Species distributions represent intraspecific genetic diversity of freshwater fish in conservation assessments. *Freshw. Biol.* **61**, 1707–1719.

14. Hoffmann AA, Sgrò CM. 2011 Climate change and evolutionary adaptation. *Nature* **470**, 479–485.

15. Carroll SP, Jorgensen PS, Kinnison MT, Bergstrom CT, Denison RF, Gluckman P, Smith TB, Strauss SY, Tabashnik BE. 2014 Applying evolutionary biology to address global challenges. *Science* **346**, 1245993–1245993.

16. Rey O, Danchin E, Mirouze M, Loot C, Blanchet S. 2016 Adaptation to global change: A transposable element–epigenetics perspective. *Trends Ecol. Evol.* **31**, 514–526.

17. Mittell EA, Nakagawa S, Hadfield JD. 2015 Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* **18**, 772–778.

18. Pritchard JK, Stephens M, Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

19. Putman AI, Carbone I. 2014 Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecol. Evol.* **4**, 4399-4428.

20. Blanchet S, Prunier JG, De Kort H. 2017 Time to go bigger: Emerging patterns in macrogenetics. *Trends Genet.* **33**, 579–580.

21. Moritz C. 1994 Defining 'Evolutionarily Significant Units' for conservation. *Trends Ecol. Evol.* **9**, 373–375.

22. Waples RS *et al.* 2001 Characterizing diversity in salmon from the Pacific Northwest. *J. Fish Biol.* **59**, 1–41.

23. Frankham R, Briscoe DA, Ballou JD. 2002 *Introduction to conservation genetics*. Cambridge, UK; New York: Cambridge University Press.

24. Neel MC, Cummings MP. 2003 Effectiveness of conservation targets in capturing genetic diversity. *Conserv. Biol.* **17**, 219–229.

25. Kottelat M, Freyhof J. 2007 *Handbook of European freshwater fishes*. Cornol, Switzerland: Publications Kottelat.

26. Poulet N, Beaulaton L, Dembski S. 2011 Time trends in fish populations in metropolitan France: insights from national monitoring data. *J. Fish Biol.* **79**, 1436–1452.

27. Paz-Vinas I, Quéméré E, Chikhi L, Loot G, Blanchet S. 2013 The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Mol. Ecol.* **22**, 3279–3291.

28. Paz-Vinas I, Comte L, Chevalier M, Dubut V, Veyssiere C, Grenouillet G, Loot G, Blanchet S. 2013 Combining genetic and demographic data for prioritizing conservation actions: insights from a threatened fish species. *Ecol. Evol.* **3**, 2696–2710.

29. Aljanabi SM, Martinez I. 1997 Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25**, 4692–4693.

30. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004 *micro-checker*: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535–538.

31. Jombart T. 2008 *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405.

32. Rousset F. 2008 *genepop'007*: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106.

33. Szpiech ZA, Jakobsson M, Rosenberg NA. 2008 *ADZE*: a rarefaction approach for counting alleles private to combinations of populations. *Bioinforma. Oxf. Engl.* **24**, 2498–2504.

34. Petit RJ, El Mousadik A, Pons O. 1998 Identifying populations for conservation on the basis of genetic markers. *Conserv. Biol.* **12**, 844–855.

35. Kalinowski ST. 2004 Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conserv. Genet.* **5**, 539–543.

36. Winter DJ. 2012 *mmod*: an R library for the calculation of population differentiation statistics. *Mol. Ecol. Resour.* **12**, 1158–1160.

37. Jost L. 2008 G(st) and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026.

38. Peterson E, Ver Hoef JMV. 2014 *STARS*: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *J. Stat. Softw.* **56**, 1–17.

39. Ver Hoef J, Peterson E, Clifford D, Shah R. 2014 *SSN*: An R Package for Spatial Statistical Modeling on Stream Networks. *J. Stat. Softw.* **56**, 1–43.

40. Ball IR, Possingham HP, Watts M. 2009 Marxan and relatives: Software for spatial conservation prioritisation. In *Spatial conservation prioritisation: Quantitative methods and computational tools* (eds A Moilanen, KA Wilson, HP Possingham), pp. 185–195. Oxford Iniversity Press, Oxford, UK.

41. Hermoso V, Kennard MJ. 2012 Uncertainty in coarse conservation assessments hinders the efficient achievement of conservation goals. *Biol. Conserv.* **147**, 52–59.

42. Finn DS, Bonada N, Múrria C, Hughes JM. 2011 Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. *J. North Am. Benthol. Soc.* **30**, 963–980.

43. Freeman LC. 1977 A Set of measures of centrality bsed on betweenness. *Sociometry* **40**, 35.

44. Estrada E, Bodin Ö. 2008 Using network centrality measures to manage landscape connectivity. *Ecol. Appl.* **18**, 1810–1825.

45. Altermatt F. 2013 Diversity in riverine metacommunities: a network perspective. *Aquat. Ecol.* **47**, 365–377.

46. Fourtune L, Paz-Vinas I, Loot G, Prunier JG, Blanchet S. 2016 Lessons from the fish: a multi-species analysis reveals common processes underlying similar species-genetic diversity correlations. *Freshw. Biol.* **61**, 1830–1845.

47. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000 Biodiversity hotspots for conservation priorities. *Nature* **403**, 853–858.

48. Fortuna MA, Albaladejo RG, Fernandez L, Aparicio A, Bascompte J. 2009 Networks of spatial genetic variation across species. *Proc. Natl. Acad. Sci.* **106**, 19044–19049.

49. Taberlet P *et al.* 2012 Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecol. Lett.* **15**, 1439–1448.

50. Paz-Vinas I, Loot G, Stevens VM, Blanchet S. 2015 Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Mol. Ecol.* **24**, 4586–4604.

51. Paz-Vinas I, Blanchet S. 2015 Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *J. Evol. Biol.* **28**, 986–994.

52. Thomaz AT, Christie MR, Knowles LL. 2016 The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution* **70**, 731–739.

53. Mimura M *et al.* 2017 Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evol. Appl.* **10**, 121–139.

54. Pauls SU *et al.* 2014 Integrating molecular tools into freshwater ecology: developments and opportunities. *Freshw. Biol.* **59**, 1559–1576.

55. Siefert A *et al.* 2015 A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecol. Lett.* **18**, 1406-1419.

Figure captions

Figure 1

Spatial distribution of observed (coloured circles) and interpolated (coloured lines) values of *AR*, *PA* and D_{est} (A, B, C respectively in the figure) for *Squalius cephalus*, *Gobio occitaniae*, *Barbatula barbatula*, and *Phoxinus phoxinus* (1, 2, 3, 4 respectively in the figure) obtained with GLMSSNs. The width of coloured lines is inversely related to the prediction standard error. The cursor on the vertical coloured scale indicates the mean value of *AR*, *PA* and D_{est}. The slope (β) of each explanatory variable (i.e. topological distance to the outlet, longitude and latitude) and its significance is also reported. N.I. indicates that the explanatory variable has not been included in the model; * p-value <0.05; ** p-value <0.01; *** p-value <0.001.

Figure 2

Percentage of irreplaceable sites identified by Marxan for conservation targets of 50, 75, 90 and 100% of the total number of alleles present in the river basin for each common species and for a pooled analysis in which all alleles from all common species were pooled together. N_{SITES} represents the number of sites included in the Marxan analyses.

Figure 3

Irreplaceable sites that have been identified for each species by Marxan for preserving 90% of the total number of alleles present in the river basin (red-filled circles).

Figure 4

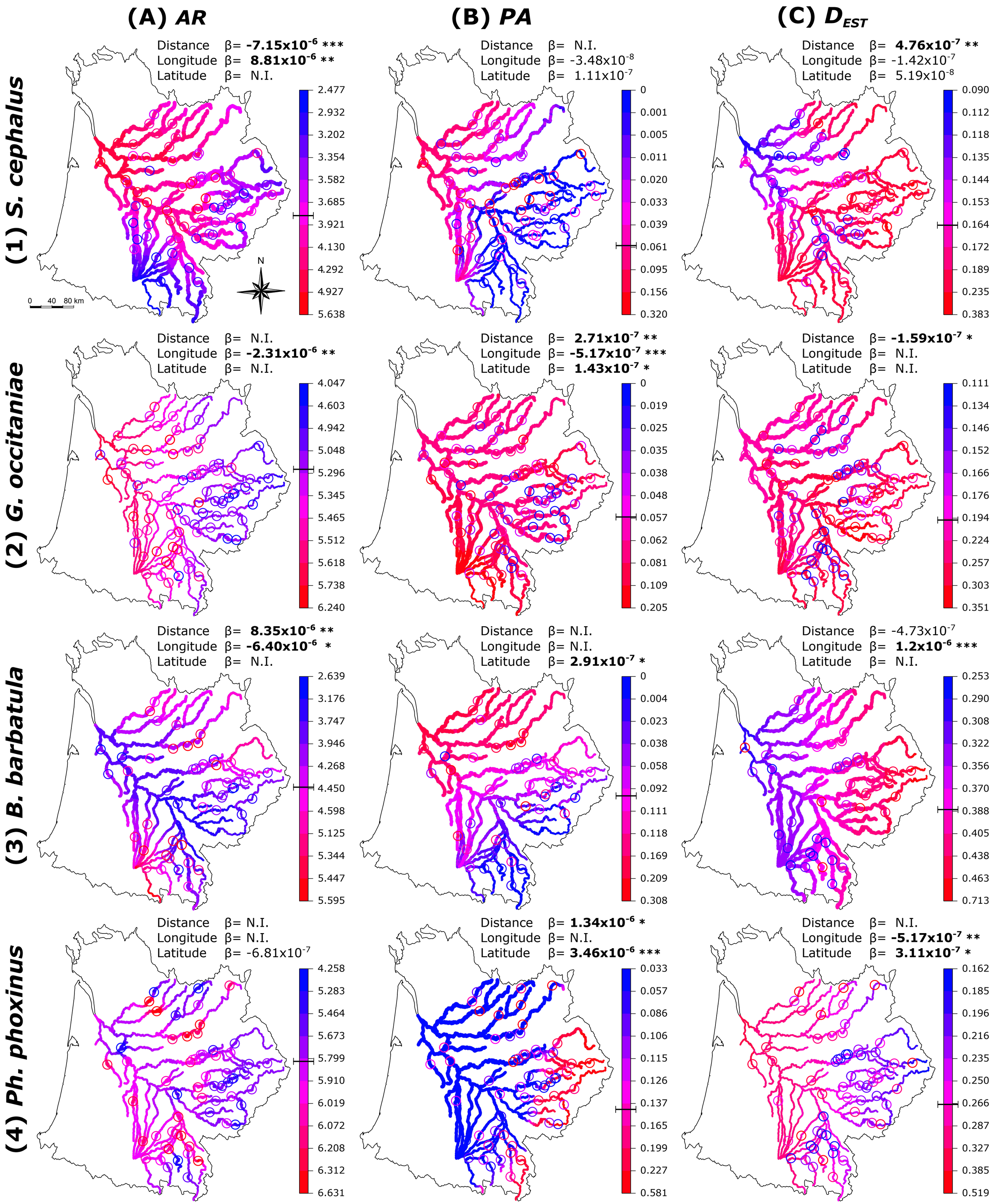
Irreplaceable sites that have been identified by Marxan for at least one (unicoloured-filled points), two (black dotted circles surrounding bicoloured points) or three (black bolded circle surrounding a tricoloured point) species, assuming a conservation target of 90% of the total number of alleles present in the river basin when considering the six species.

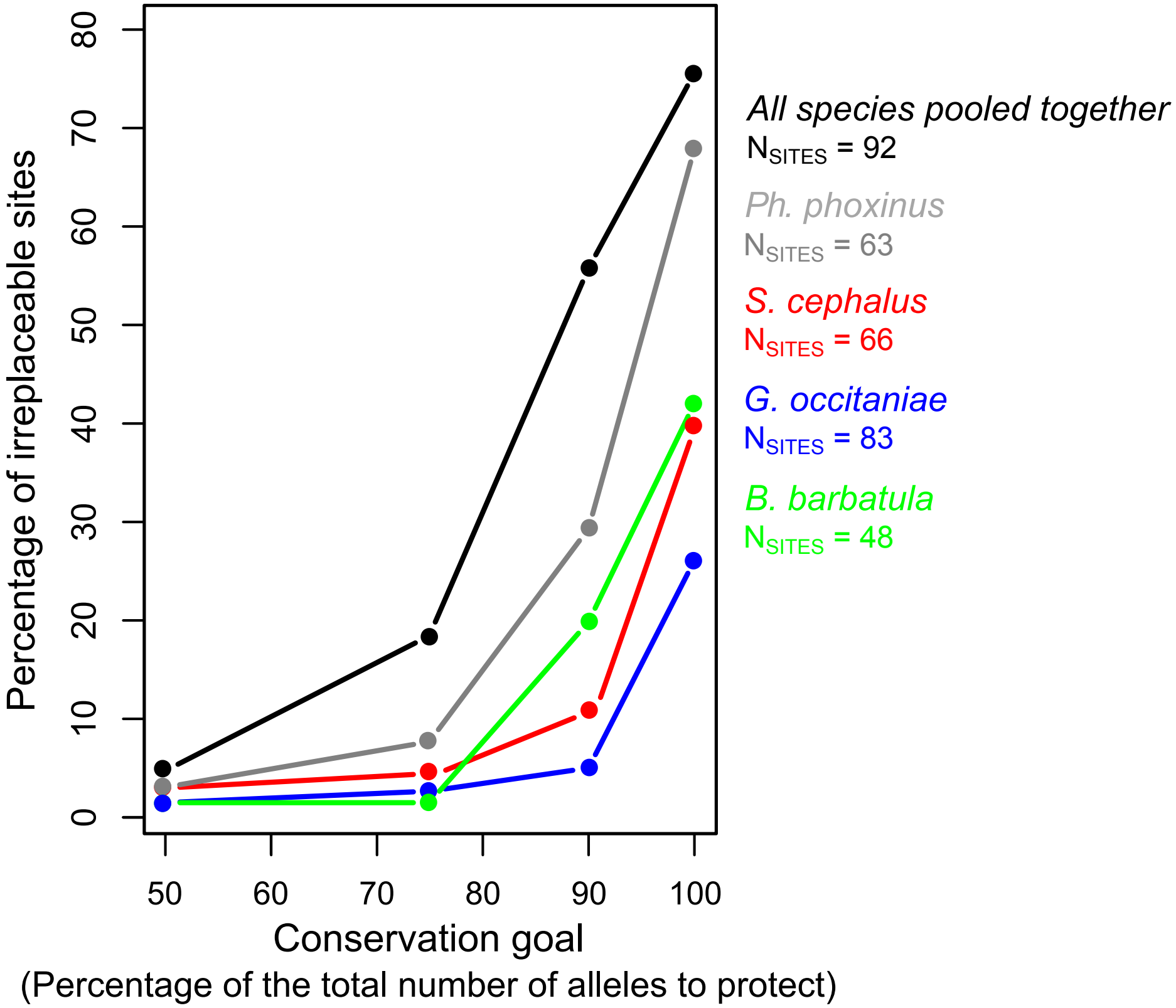
Figure 5

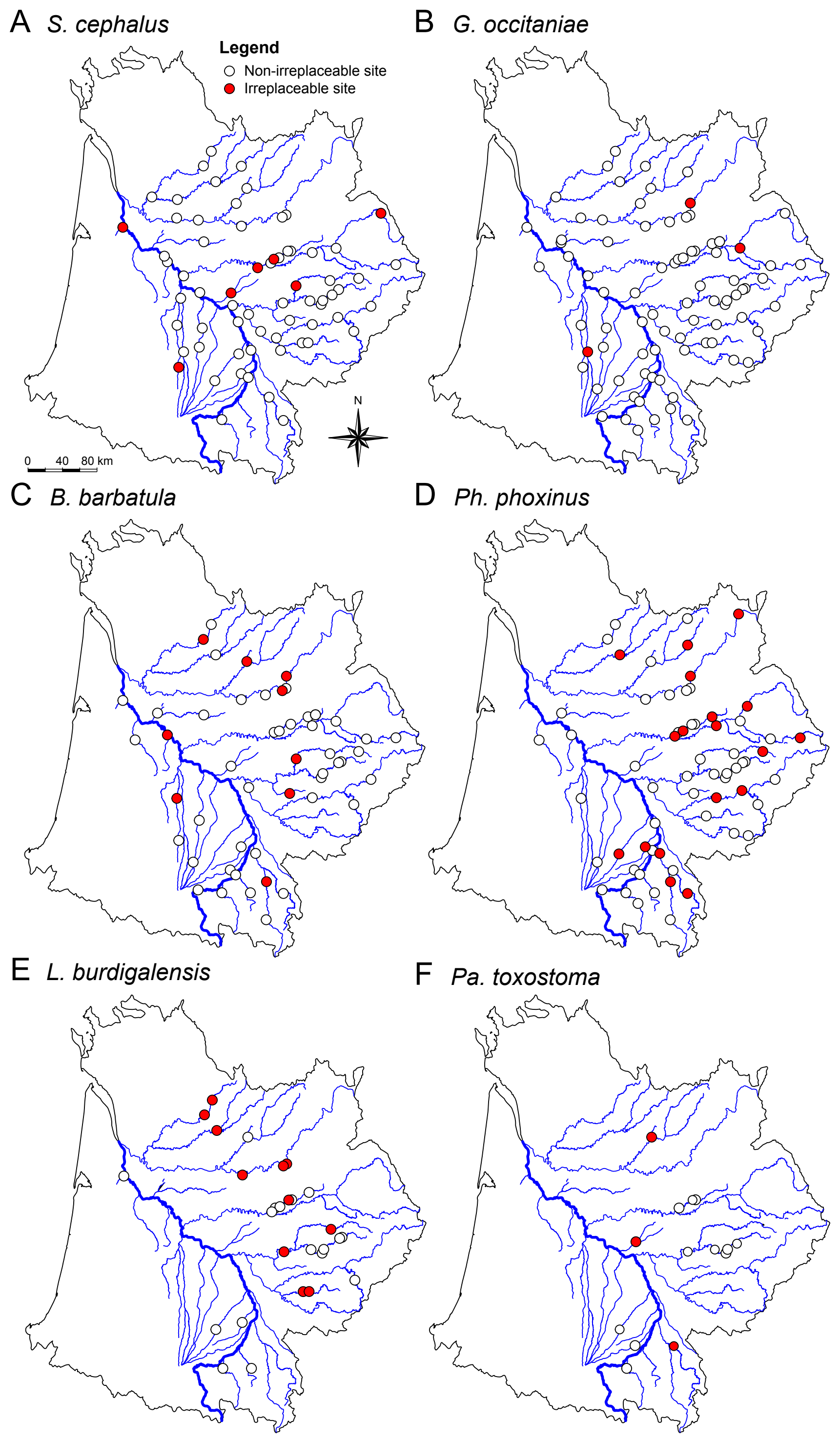
Mean percentage of the total number of alleles observed at the riverscape scale for all the six species that are secured by (i) a cumulative number of randomly selected sampling sites (blue line; average accumulation curve obtained under 1000 permutations of the sampling sites sites; grey vertical lines: standard error), (ii) irreplaceable sites identified using SCP for each of the six species independently (red dots), (iii) irreplaceable sites identified using SCP for different sets of species (red squares), and (iv) sites identified on the basis of freshwater fish species co-occurrence data (black dots).

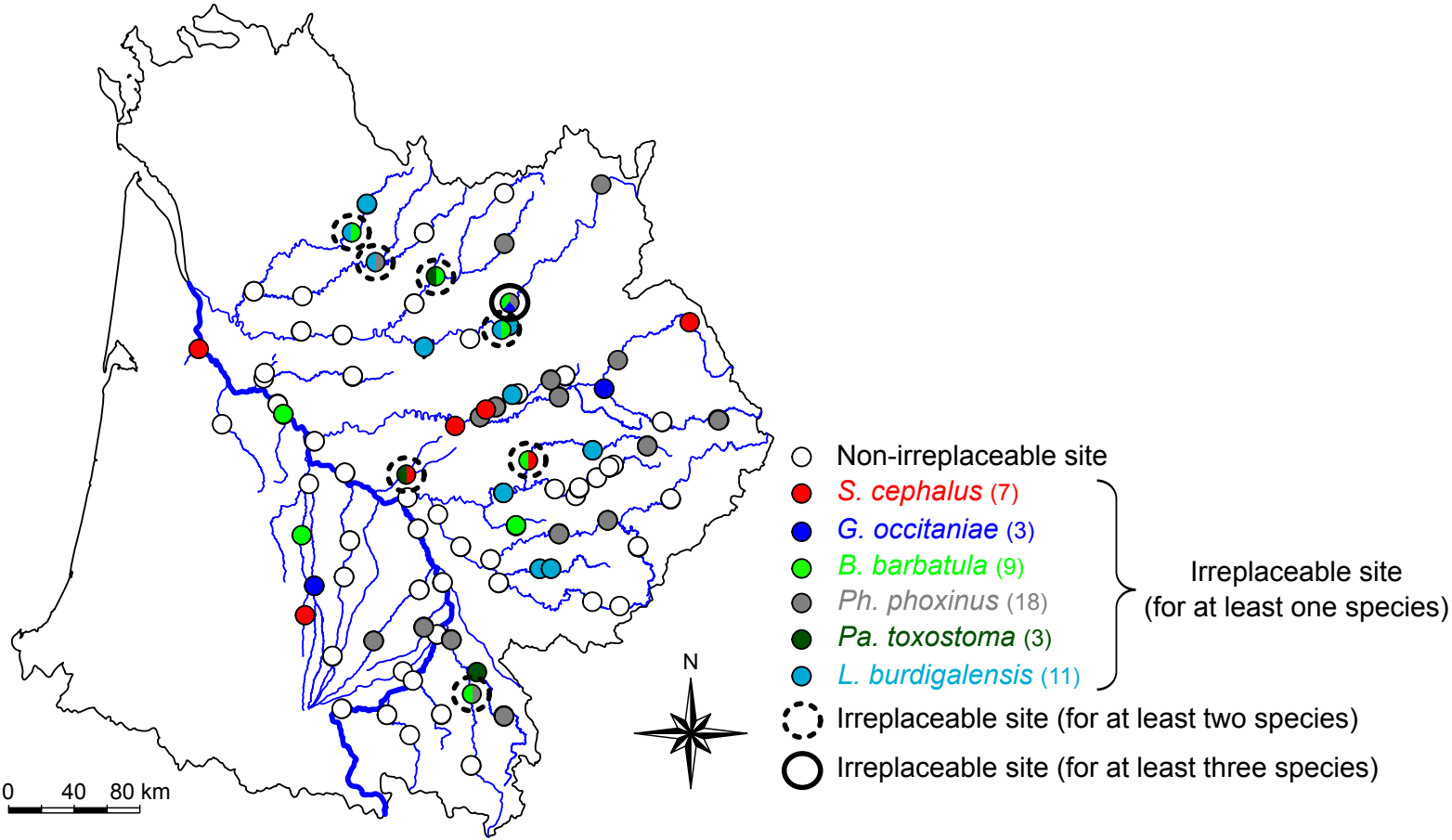
TABLE 1: Table reporting for each species the percentages of their total number of alleles that are secured by the sampling sites having been identified as irreplaceable sites for the other species (considering the 90% conservation target).

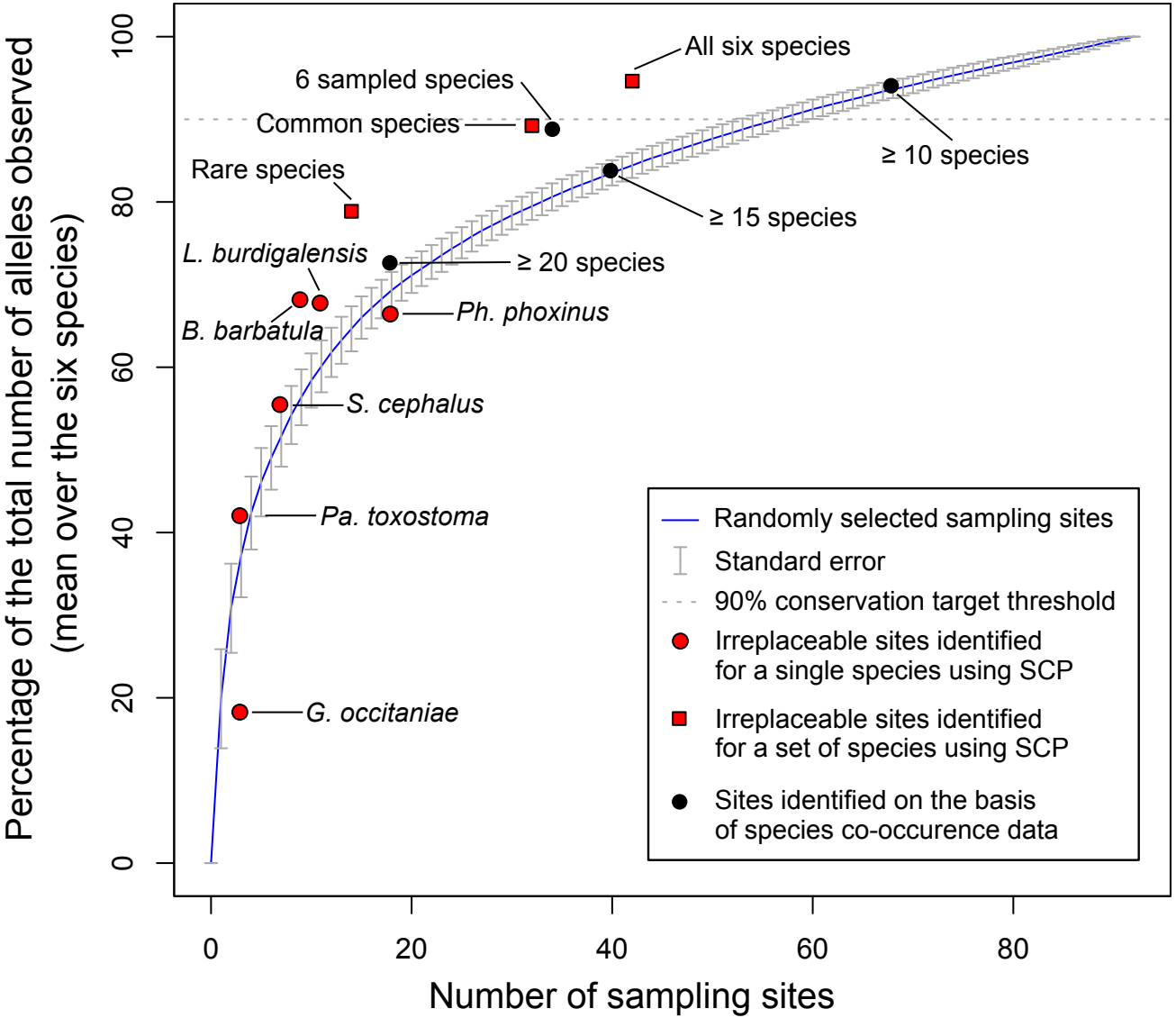
Proportion of alleles from (below) that fell within irreplaceable sites from (right):	*Squalius cephalus* *n=7*	*Gobio occitaniae* *n=3*	*Barbatula barbatula* *n=9*	*Phoxinus phoxinus* *n=18*	*Leuciscus burdigalensis* *n=11*	*Parachondrostoma toxostoma* *n=3*	Common species *n=32*	Rare species *n=14*
Squalius cephalus	*90.00*	17.66	61.52	56.52	61.60	39.73	—	65.18
Gobio occitaniae	65.60	*90.00*	66.76	79.58	70.27	50.97	—	73.40
Barbatula barbatula	49.79	31.65	*90.00*	79.57	76.78	38.95	—	79.40
Phoxinus phoxinus	48.25	42.60	68.43	*90.00*	63.49	44.61	—	68.88
Leuciscus burdigalensis	44.73	0	66.13	67.09	*90.00*	36.42	81.79	—
Parachondrostoma toxostoma	78.55	0	78.55	50.00	67.14	*90.00*	90.00	—
Average over all other species	*55.58*	*18.38*	*68.28*	*66.55*	*67.86*	*42.14*		


(A) AR
(B) PA
(C) D_EST
(1) S. cephalus
(2) G. occitaniae
(3) B. barbatula
(4) Ph. phoxinus
Distance β= -7.15x10^-6 ***
Longitude β= 8.81x10^-6 **
Latitude β= N.I.
2.477 2.932 3.202 3.354 3.582 3.685 3.921 4.130 4.292 4.927 5.638
Distance β= N.I.
Longitude β= -3.48x10^-8
Latitude β= 1.11x10^-7
0 0.001 0.005 0.011 0.020 0.033 0.039 0.061 0.095 0.156 0.320
Distance β= 4.76x10^-7 **
Longitude β= -1.42x10^-7
Latitude β= 5.19x10^-8
0.090 0.112 0.118 0.135 0.144 0.153 0.164 0.172 0.189 0.235 0.383
Distance β= N.I.
Longitude β= -2.31x10^-6 **
Latitude β= N.I.
4.047 4.603 4.942 5.048 5.296 5.345 5.465 5.512 5.618 5.738 6.240
Distance β= 2.71x10^-7 **
Longitude β= -5.17x10^-7 ***
Latitude β= 1.43x10^-7 *
0 0.019 0.025 0.035 0.038 0.048 0.057 0.062 0.081 0.109 0.205
Distance β= -1.59x10^-7 *
Longitude β= N.I.
Latitude β= N.I.
0.111 0.134 0.146 0.152 0.166 0.176 0.194 0.224 0.257 0.303 0.351
Distance β= 8.35x10^-6 **
Longitude β= -6.40x10^-6 *
Latitude β= N.I.
2.639 3.176 3.747 3.946 4.268 4.450 4.598 5.125 5.344 5.447 5.595
Distance β= N.I.
Longitude β= N.I.
Latitude β= 2.91x10^-7 *
0 0.004 0.023 0.038 0.058 0.092 0.111 0.118 0.169 0.209 0.308
Distance β= -4.73x10^-7
Longitude β= 1.2x10^-6 ***
Latitude β= N.I.
0.253 0.290 0.308 0.322 0.356 0.370 0.388 0.405 0.438 0.463 0.713
Distance β= N.I.
Longitude β= N.I.
Latitude β= -6.81x10^-7
4.258 5.283 5.464 5.673 5.799 5.910 6.019 6.072 6.208 6.312 6.631
Distance β= 1.34x10^-6 *
Longitude β= N.I.
Latitude β= 3.46x10^-6 ***
0.033 0.057 0.086 0.106 0.115 0.126 0.137 0.165 0.199 0.227 0.581
Distance β= N.I.
Longitude β= -5.17x10^-7 **
Latitude β= 3.11x10^-7 *
0.162 0.185 0.196 0.216 0.235 0.250 0.266 0.287 0.327 0.385 0.519
N
0 40 80 km


Percentage of irreplaceable sites
0
10
20
30
40
50
60
70
80
50
60
70
80
90
100
Conservation goal
(Percentage of the total number of alleles to protect)
All species pooled together
N SITES = 92
Ph. phoxinus
N SITES = 63
S. cephalus
N SITES = 66
G. occitaniae
N SITES = 83
B. barbatula
N SITES = 48


A S. cephalus
Legend
Non-irreplaceable site
Irreplaceable site
N
0 40 80 km
B G. occitaniae
C B. barbatula
D Ph. phoxinus
E L. burdigalensis
F Pa. toxostoma


Non-irreplaceable site
S. cephalus (7)
G. occitaniae (3)
B. barbatula (9)
Ph. phoxinus (18)
Pa. toxostoma (3)
L. burdigalensis (11)
Irreplaceable site
(for at least one species)
Irreplaceable site (for at least two species)
Irreplaceable site (for at least three species)
N
0 40 80 km


Percentage of the total number of alleles observed
(mean over the six species)
Number of sampling sites
0
20
40
60
80
100
All six species
6 sampled species
Common species
Rare species
≥ 10 species
≥ 15 species
≥ 20 species
L. burdigalensis
B. barbatula
Ph. phoxinus
S. cephalus
Pa. toxostoma
G. occitaniae
Randomly selected sampling sites
Standard error
90% conservation target threshold
Irreplaceable sites identified for a single species using SCP
Irreplaceable sites identified for a set of species using SCP
Sites identified on the basis of species co-occurence data